

Adempimenti per l'avvio di studi osservazionali retrospettivi in medicina di laboratorio

Fabio Del Ben¹, Alice Pisapia², Andrea Padoan³, Roberta Rolla⁴, Tommaso Trenti⁵, Giuseppe Banfi^{6,7}, Anna Carobene⁸,
Rossella Tomaiuolo^{6,7}

¹IRCCS Centro di Riferimento Oncologico di Aviano (Pordenone)

²Università degli Studi dell'Insubria (Varese)

³Università di Padova

⁴Università del Piemonte Orientale (Novara)

⁵Ospedale Civile Sant'Agostino-Estense (Baggiovara, Modena)

⁶Università Vita-Salute San Raffaele (Milano)

⁷IRCCS Ospedale Galeazzi-Sant'Ambrogio (Milano)

⁸IRCCS Ospedale San Raffaele (Milano)

ABSTRACT

Requirements for Initiating Retrospective Observational Studies in Laboratory Medicine

In recent years, the increasing digitalization of health technologies and systems for collecting and storing health data has allowed for the creation of value from data and its use. However, this has led to the need to implement risk management regarding the protection and loss of health data. Observational studies, by nature, do not pose direct additional risks to the health of patients. However, the privacy of patients' needs to be preserved. This document, by presenting the obligations introduced by the current Regulation and delving into its impacts on governance, security, and management of health data, aims to provide technical-organizational measures to support the use of health data for research purposes in retrospective observational studies, particularly in laboratory medicine. It should be noted that this document is merely illustrative with reference to the cases analyzed and does not intend in any way to eliminate the evaluations of the Data Protection Officer (DPO) of the involved structures, as well as the necessary and timely risk assessment reserved for the Data Controller (DC). The document therefore aims to support the researcher in the task of providing the DPO and the DC with all the preliminary information regarding the study, in order to simplify and speed up the evaluation process. Accordingly, the document has been structured so that the researcher can benefit not only from information on the current legislation but also from informal tools such as dialogue windows with the lawyer and the analysis of real-world case studies.

Parole chiave: medicina di laboratorio, etica, studi osservazionali

PREMESSA

Negli ultimi anni, la crescente digitalizzazione delle tecnologie sanitarie e dei sistemi di raccolta e di archiviazione dei dati sanitari ha permesso di creare valore a partire dai dati e dal loro utilizzo (1,2). Tale evoluzione, tuttavia, ha portato alla necessità di implementare la gestione dei rischi relativi alla tutela e alla dispersione di dati sanitari (3).

Questo documento, presentando gli adempimenti introdotti dal Regolamento vigente (4) e approfondendone gli impatti relativi a governance, sicurezza e gestione dei dati sanitari, si propone di fornire delle misure tecnico-

organizzative volte a sostenere l'utilizzo dei dati sanitari a scopo di ricerca negli studi retrospettivi osservazionali.

Si precisa che il presente documento ha una portata meramente esemplificativa con riferimento ai casi analizzati e non intende in alcun modo eliminare le valutazioni del Data Protection Officer (DPO) delle strutture coinvolte, così come la necessaria e puntuale valutazione del rischio riservata al Data Controller (DC). Tali valutazioni dovranno infatti essere effettuate sul singolo caso concreto prendendo in considerazione tutti gli elementi del trattamento dei dati: tipologia dati, base giuridica, eventuale consenso sottoscritto, finalità del trattamento stesso e mitigazione dei rischi.

Corrispondenza a: Fabio Del Ben, IRCCS Centro di Riferimento Oncologico di Aviano (PN), Email: fabio.delben@cro.it

Ricevuto: 14.11.2023

Revisionato: 02.12.2023

Accettato: 06.12.2023

Pubblicato online: 08.01.2024

DOI: 10.19186/BC_2023.094

Come chiarito anche da recente giurisprudenza del Tribunale dell'Unione (UE) (5), è essenziale valutare l'anonimizzazione o la pseudonimizzazione dei dati nei contesti specifici del caso concreto essendo necessaria una valutazione puntuale del rischio privacy per l'applicazione del Regolamento 679/2016 (4).

Il documento si propone quindi di supportare il ricercatore nel compito di fornire al DPO e al titolare del trattamento, le informazioni in riferimento alla tipologia dei dati utilizzati per lo studio al fine di semplificare e velocizzare il processo valutativo.

Pertanto, il documento è stato strutturato in modo che il ricercatore possa usufruire non solo di informazioni sulla normativa vigente, ma anche di strumenti informali quali le finestre di dialogo con il giurista (vedi la Figura 1 come esempio) e l'analisi di casi reali.

Il documento è stato redatto da un gruppo multidisciplinare costituito da ricercatori, accademici e ospedalieri, esperti nel campo della medicina di laboratorio e da una giurista specializzata nel campo della ricerca medico-scientifica. Il gruppo si è spontaneamente costituito attorno ad un problema di natura etica, particolarmente rilevante per la professione ed ha trovato un consenso sulle posizioni esposte.

INTRODUZIONE

Gli studi osservazionali non comportano rischi aggiuntivi per i pazienti e dunque necessitano di una disciplina specifica, diversa da quella prevista per gli studi interventistici, che ne agevoli e ne promuova la realizzazione (6).

La conduzione di uno studio osservazionale, al pari degli altri studi clinici, presuppone sempre il coinvolgimento, da parte del promotore, del competente Comitato Etico, con modalità che si differenziano a seconda che lo studio sia prospettico (valutazione degli effetti di un intervento attraverso dati raccolti durante lo studio) o retrospettivo (valutazione di eventi accaduti in un periodo precedente al disegno dello studio, attraverso dati raccolti prima che lo studio iniziasse).

Per gli studi prospettici, mentre le Linee guida del 2008 (7) prevedevano che i Comitati Etici diversi dal comitato

coordinatore, potessero inviare a quest'ultimo eventuali osservazioni e comunicare "formale o motivato rifiuto per la partecipazione allo studio", il decreto del 30 novembre 2021 (8), stabilisce che non è necessaria l'approvazione di tutti i Comitati Etici dei centri partecipanti, semplificando la procedura.

Riguardo invece l'avvio di studi osservazionali retrospettivi, le Linee Guida del 2008 prescrivevano la notifica ad opera del promotore a tutti i Comitati Etici dei centri coinvolti, alla quale si accompagnava un meccanismo di silenzio-assenso che consentiva di iniziare lo studio decorsi 60 giorni dal tale adempimento. Il decreto del 30 novembre 2021 (8) non si sofferma su questo aspetto, limitandosi a prendere in considerazione il parere unico nel caso di studi prospettici multicentrici.

Gli studi osservazionali sono una categoria molto ampia ed eterogenea, sia per obiettivo che per oggetto. Questi studi sono caratterizzati dall'assenza di intervento attivo da parte dei ricercatori (ad esempio, il ricercatore non determina l'assegnazione dei soggetti ai diversi gruppi di studio), e dalla pura registrazione di quello che avviene nella realtà. Altra caratteristica degli studi osservazionali è che l'oggetto di indagine può essere estremamente vario, costituito da persone sane o malate, farmaci, dispositivi medici, procedure diagnostiche, biomarcatori, fattori prognostici e predittivi, geni coinvolti nelle malattie ereditarie, rischi professionali, cambiamenti della pratica clinica, appropriatezza delle procedure, ed altro. Inoltre, è caratteristico il trattamento dei dati sanitari dei soggetti inclusi nello studio, in particolare modo negli studi retrospettivi, considerando il tempo trascorso dagli eventi oggetto di analisi. Nel caso di studi osservazionali retrospettivi, secondo l'articolo 110 del decreto legislativo 196/2003 (9), quando informare gli interessati risulti impossibile o implichi per il titolare uno sforzo sproporzionato, il trattamento può prescindere dal consenso se il programma di ricerca, una volta ottenuto il parere favorevole del competente Comitato Etico, è sottoposto alla consultazione del Garante ai sensi dell'articolo 36 del General Data Protection Regulation (GDPR) (4) e ottiene il relativo parere favorevole (Figura 1).

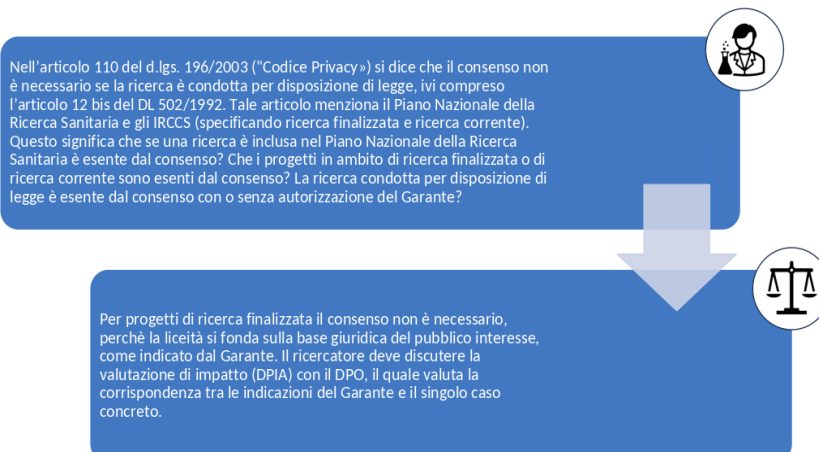


Figura 1

Esempio di domanda del ricercatore e risposta del giurista: consenso informato e ricerca condotta per disposizione di legge

Tuttavia, l'applicazione dell'articolo 110 richiede un percorso lungo e complesso per cui, al fine di incentivare la conduzione di studi osservazionali, sarebbe certamente auspicabile una semplificazione degli adempimenti necessari, ad esempio consentendo l'utilizzo generalizzato di basi giuridiche alternative al consenso senza dovere dipendere, volta per volta, da singoli pareri o autorizzazioni del Garante. In questo senso, anche il documento del Centro di Coordinamento Nazionale dei Comitati Etici (10) è a favore di un cambiamento nell'attuale approccio, auspicando la rimozione degli ostacoli all'utilizzo e al riutilizzo dei dati di ricerca, anche attraverso l'impiego del legittimo interesse quale base giuridica per il trattamento dei dati, accompagnato naturalmente dall'adozione di adeguate misure di sicurezza per la tutela degli interessati.

Esistono inoltre, per lo stesso articolo 110, casi particolari di ricerca definita da disposizioni di legge o regolamenti, come ad esempio la ricerca finalizzata negli Istituti di Ricovero e Cura a Carattere Scientifico (IRCCS), per i quali il consenso dell'interessato non è necessario, in quanto la liceità delle attività poggia sulla base giuridica del pubblico interesse.

Ciò nonostante, come noto, in Italia l'utilizzo sistematico di basi giuridiche alternative al consenso è ostacolato da una normativa nazionale che tutt'ora vede nel consenso la principale fonte di legittimazione del trattamento, e consente di prescindervi solamente nel rispetto di procedure e adempimenti spesso molto difficili da percorrere in concreto.

Tra i professionisti della medicina di laboratorio è nata dunque l'esigenza di sviluppare un documento che fornisca le basi per un inquadramento dei processi e delle nozioni giuridiche relative a studi osservazionali retrospettivi basati sull'analisi dei dati di laboratorio. Pertanto, dalla collaborazione interdisciplinare tra professionisti di laboratorio e giuristi, è stato preparato questo documento che, dopo una panoramica sull'opportunità e i rischi dell'utilizzo dei dati di laboratorio, affronta, con un taglio rigoroso normativamente ma pratico e fruibile, argomenti essenziali come la protezione dei dati personali, e il GDPR nell'ambito della ricerca medica epidemiologica. Nella parte finale del documento, sono riportati alcuni esempi di situazioni tipiche con il relativo commento giuridico.

IL DATO DI LABORATORIO: OPPORTUNITÀ E RISCHI

La medicina di laboratorio produce ogni giorno una grande quantità di dati, tra i quali bisogna distinguere i dati relativi alle analisi di laboratorio di base (ad esempio la chimica clinica, l'ematologia o gli esami di coagulazione), i dati "omici" (quelli provenienti da piattaforme di genomica, metabolomica e proteomica) e i dati non relativi al paziente (ad esempio, i dati di manutenzione strumentale, calibrazione, controlli di qualità esterni ed interni) (11). Arricchiti da dati clinici, i dati di medicina di laboratorio consentono di acquisire nuove conoscenze fisiopatologiche, sono in grado di migliorare gli esiti del paziente anche attraverso terapie personalizzate (12), possono essere utilizzati per sviluppare intervalli

di riferimento a fini diagnostici, o per valutazioni di Health Technology Assessment (HTA) di tecnologie sanitarie (13). Probabilmente, il campo di applicazione più interessante per la medicina di laboratorio è la diagnostica predittiva e preventiva. Con il supporto dei dati di laboratorio, è possibile calcolare le probabilità per una varietà di eventi correlati al paziente e avviare contromisure. Tale approccio si applica potenzialmente a tutte le fasi di cura del paziente, dalla prevenzione, alla diagnosi, al trattamento, al monitoraggio della terapia, e di eventuali recidive (14).

Un altro aspetto di grande importanza è il contributo della medicina di laboratorio ai progetti di ricerca: nella stragrande maggioranza degli studi clinici, i dati di laboratorio svolgono un ruolo estremamente importante (ad esempio, come variabili di risultato, come valori di riferimento, come indicatori di qualità e conformità). Oltre alla diagnostica di laboratorio stessa, i dati di laboratorio trovano applicazione nel campo della gestione della qualità, nell'ottimizzazione degli aspetti amministrativi e orientati ai processi, come l'ottimizzazione della fase pre-analitica, analitica e post-analitica.

La medicina di laboratorio è una disciplina basata sui dati: oggi più che mai, con l'avvento delle tecnologie multiparametriche, la disponibilità dei dati rappresenta la base per le moderne tecnologie, in particolare per il machine learning o l'intelligenza artificiale, che possono portare vantaggi non solo per gli esiti di salute, ma anche per la medicina di laboratorio stessa come disciplina medica accademica e basata sulla ricerca (15,16). Sebbene siano necessari molti passaggi per la trasformazione dei dati della medicina di laboratorio in Big Data, bisogna affrontare la sfida di utilizzare i database di laboratorio non più come voluminosi contenitori di caratteri alfanumerici ma come miniera di informazioni per intenti di ricerca e sviluppo mirati al miglioramento della di salute, implementando approcci di data mining.

Esistono molti tipi diversi di database, scelti ed organizzati secondo il modo in cui si intendono utilizzare i dati. In genere, i database di laboratorio contengono dati identificativi (ID), dati quasi identificativi (QID) e dati sensibili (Tabella 1) (17).

L'utilizzo di dati di laboratorio non ha solo vantaggi, ma è anche associato a un numero considerevole di rischi: come per tutti i dati sanitari, anche per i dati di laboratorio esiste il rischio di fughe di dati, soprattutto se i dati sono accessibili dall'esterno. Inoltre, i dati di laboratorio possono anche essere copiati facilmente e rapidamente, con un notevole rischio di duplicazione non autorizzata. È essenziale, quindi, proteggere le informazioni personali e sensibili, per rendere difficile identificare gli individui a cui i dati si riferiscono. L'anonimizzazione è una procedura che impedisce di collegare le informazioni disponibili ad una persona fisica identificata o identificabile; pertanto, i dati personali resi sufficientemente anonimi non consentono più l'identificazione dell'interessato. La pseudonimizzazione, invece, consente che i dati possano essere attribuiti a una persona fisica mediante l'utilizzo di informazioni ulteriori. Tali dati dovrebbero essere considerati informazioni su una persona fisica identificabile.

Tabella 1*Definizione dei diversi tipi di dati*

Tipo di dati	Definizione	Esempi
Identificativi (ID)	Dati che identificano direttamente una persona	Generalità, codice fiscale, indirizzo, numero di telefono
Quasi-identificativi (quasi-identifier, QID)	Dati che di per se' non identificano la persona, ma che combinati tra di loro possono ricondurre all'identificazione della persona.	Sesso, data di nascita, provincia di nascita
Sensibili	Dati che riguardano aspetti confidenziali della persona	Patologia in atto o pregressa, familiarità per patologia, terapia farmacologica

Per stabilire l'identificabilità di una persona è necessario considerare tutti i mezzi di cui il titolare del trattamento può ragionevolmente avvalersi per identificare direttamente la persona fisica. Per accertare la ragionevole probabilità di utilizzo dei mezzi per identificare la persona fisica, si dovrebbe prendere in considerazione l'insieme dei fattori obiettivi, tra cui i costi e il tempo necessario per l'identificazione, tenendo conto sia delle tecnologie disponibili al momento del trattamento, sia degli sviluppi tecnologici. Come precisato dalla sentenza del Tribunale dell'UE dell'aprile 2023 (5), è necessario verificare in concreto se i soggetti coinvolti nel trattamento dei dati personali abbiano ragionevole possibilità di reidentificare gli individui a cui i dati si riferiscono anche grazie ad informazioni aggiuntive o se, al contrario, nella pratica essi non dispongano di mezzi legali per operare tale procedura. Non vige quindi una presunzione assoluta, ma solo una valutazione in concreto dei rischi di reidentificazione. Nella Tabella 2 sono riportate alcune delle tecniche comunemente utilizzate (17), singolarmente o in combinazione, a seconda delle esigenze di privacy e delle caratteristiche del database (18-20). È importante notare che l'efficacia di queste tecniche dipende dalla loro corretta implementazione e dall'equilibrio tra la protezione della privacy e la conservazione dell'utilità dei dati.

L'applicazione di queste tecniche farà acquisire al database specifiche proprietà, che ne indicano il grado di anonimizzazione (18). Alcuni esempi di proprietà di un database, fra i più comuni, sono descritti in Tabella 3. Ad esempio, con soppressione e generalizzazione si può ottenere un database $k=3$ -anonimo, ovvero un database in cui per ogni individuo vi sono almeno altri due individui con le medesime variabili sensibili, in modo da rendere l'individuo non identificabile.

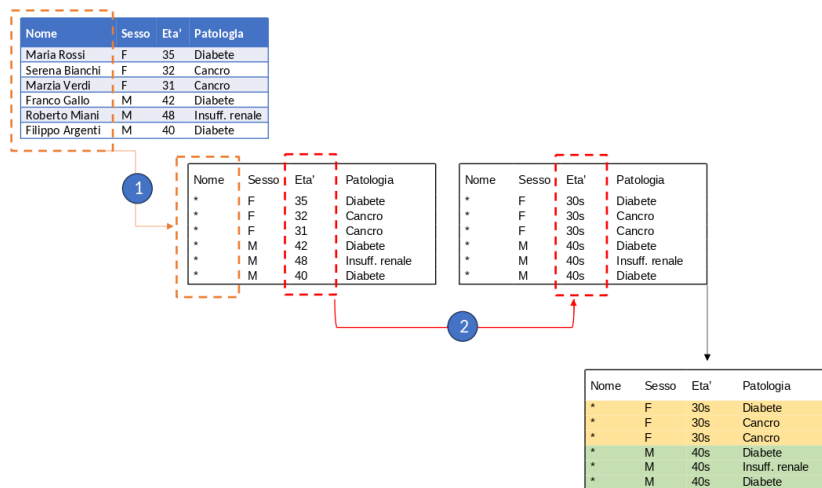
Nella Figura 2 è schematizzato il processo di trasformazione di un database. In questo processo sono state applicate Soppressione al nome, e Generalizzazione all'età. Sono state realizzate due classi di equivalenza con 3 equivalenti ciascuna ($k=3$). Ogni classe di equivalenza ha almeno 2 dati sensibili unici ($l=2$), ovvero Diabete e Cancro per il primo gruppo, Diabete e Insufficienza renale per il secondo gruppo. Il database soddisfa i requisiti di anonimità $K=3$ e $l=2$, e non è pertanto facilmente attaccabile da tentativi di ricondurre il dato all'individuo. Tenuto fermo che i dati identificativi vanno sempre rimossi, questi argomenti riguardano database in cui sono pubblicati i QID. Se i QID non sono utilizzati o pubblicati, i requisiti di anonimizzazione si semplificano molto.

Tabella 2*Tecniche di anonimizzazione*

Tecnica di anonimizzazione	Modalità di operazione
Mascheramento dei dati (Data Masking)	Sostituzione di dati sensibili con dati fittizi o modificati, mantenendo una struttura simile ai dati originali
Pseudonimizzazione	Sostituzione degli identificatori diretti (ad esempio nomi o indirizzi e-mail) con pseudonimi univoci per garantire che i dati non possano essere associati facilmente agli individui a cui si riferiscono
Distorsione dei dati (Data Perturbation)	Modificazione casuale dei dati (ad esempio aggiunta di rumore), per rendere difficile identificare gli individui a cui i dati si riferiscono, pur mantenendo le caratteristiche generali dei dati
Generalizzazione	Sostituzione dei dati più specifici con dati più generali (ad esempio, aggregare le età in fasce d'età o convertire indirizzi specifici in codici postali) per ridurre la granularità dei dati, rendendo difficile l'identificazione dei singoli individui
Soppressione	Rimozione completa di determinate informazioni dal database, in modo che non possano essere utilizzate per identificare gli individui

Tabella 3*Proprietà di un database relative al grado di anomizzazione*

Proprietà	Significato
K-Anonymity (21)	Per ogni individuo, nel database, sono presenti almeno altri k individui con gli stessi QID. Tutti gli individui con gli stessi QID costituiscono una classe di equivalenza. Pertanto, individui appartenenti alla stessa classe di equivalenza non saranno identificabili
L-Diversity (22)	Estensione della k-anonizzazione, motivata dal fatto che quest'ultima non garantisce la privacy se una stessa classe di equivalenza non ha sufficiente diversità, ad esempio se tutti gli individui di una stessa classe di equivalenza hanno la stessa patologia. Un dataset ha L-diversity se per ogni insieme di righe con valori unici QID esistono almeno L valori distinti per ciascun dato sensibile.
T-Closeness (23-24)	Ulteriore estensione della due precedenti proprietà, motivata da limitazioni delle due tecniche. Per essere definita t-close, la distribuzione dei dati sensibili all'interno di ciascuna classe di equivalenza deve essere simile alla distribuzione nel database originale.

**Figura 2**

Schematizzazione del processo di trasformazione di un database contenente dati identificativi (ID), semi-identificativi (QID) e sensibili in un database k-anonimo con $k=3$, e l-diversità=2. Nel passaggio 1, il nome (ID) è soppresso. Nel passaggio 2, l'età (QID) è generalizzata. Il database risultante contiene due classi di equivalenza (una in giallo, l'altra in verde). All'interno di ciascuna classe vi sono almeno 3 equivalenti ($k=3$) e 2 dati sensibili unici (2 tipi di patologia).

Definizioni applicabili in materia di protezione dei dati personali

Il regolamento europeo 679/2016 (GDPR) (4) stabilisce norme relative alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché norme relative alla libera circolazione di tali dati nel mercato interno dell'Unione. I dati personali devono essere trattati in modo lecito, corretto e trasparente nei confronti dell'interessato. Inoltre, i dati personali devono essere raccolti per finalità determinate, esplicite e legittime, e successivamente trattati in un modo che sia compatibile con tali finalità. Nella Tabella 4 sono riportate le tipologie di dati.

Il trattamento è lecito solo se è presente almeno una delle seguenti basi giuridiche (Figura 3):

- l'interessato ha espresso il consenso al trattamento dei

propri dati personali per una o più specifiche finalità;

- il trattamento è necessario all'esecuzione di un contratto di cui l'interessato è parte, se è necessario per adempiere un obbligo legale (come per esempio obblighi fiscali);
- il trattamento è necessario per la salvaguardia degli interessi vitali dell'interessato o di un'altra persona fisica;
- il trattamento è necessario per l'esecuzione di un compito di interesse pubblico;
- il trattamento è necessario per il perseguimento del legittimo interesse del titolare del trattamento o di terzi, a condizione che non prevalgano gli interessi o i diritti e le libertà fondamentali dell'interessato che richiedono la protezione dei dati personali.

Nella Tabella 5 sono riportate le definizioni relative al trattamento e ai soggetti coinvolti.

Tabella 4
Tipologie di dati

Tipo di dato	Definizione
Personale	Qualsiasi informazione riguardante una persona fisica identificata o identificabile (interessato). Si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come il nome, un numero di identificazione (numero paziente), dati relativi all'ubicazione, un identificativo online o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale.
Genetico	Dati personali relativi alle caratteristiche genetiche ereditarie o acquisite di una persona fisica che forniscono informazioni univoche sulla fisiologia o sulla salute di detta persona fisica, e che risultano in particolare dall'analisi di un campione biologico della persona fisica in questione.
Biometrico	Dati personali ottenuti da un trattamento tecnico specifico relativi alle caratteristiche fisiche, fisiologiche o comportamentali di una persona fisica che ne consentono o confermano l'identificazione univoca, quali per esempio l'immagine facciale.
Sanitario	Dati personali attinenti alla salute fisica o mentale di una persona fisica, compresa la prestazione di servizi di assistenza sanitaria, che rivelano informazioni relative al suo stato di salute. Tali dati sono da annoverarsi nelle categorie dei dati ex articolo 9 del GDPR ovvero afferenti alle "categorie particolari di dati" (già definiti dati sensibili).

GDPR, General Data Protection Regulation

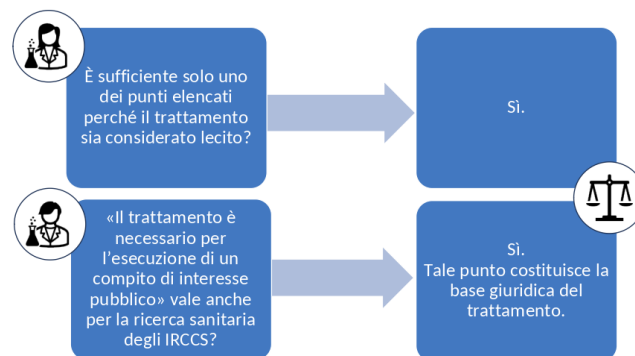


Figura 3
Domande del ricercatore e risposta del giurista relative alle basi giuridiche del trattamento.

Tabella 5
Definizioni relative al trattamento dei dati e ai soggetti coinvolti.

Nome	Definizione
Trattamento	Qualsiasi operazione o insieme di operazioni, compiute con o senza l'ausilio di processi automatizzati e applicate a dati personali o insiemi di dati personali, come la raccolta, la registrazione, l'organizzazione, la strutturazione, la conservazione, l'adattamento o la modifica, l'estrazione, la consultazione, l'uso, la comunicazione mediante trasmissione, diffusione o qualsiasi altra forma di messa a disposizione, il raffronto o l'interconnessione, la limitazione, la cancellazione o la distruzione.
Profilazione	Qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di tali dati personali per valutare determinati aspetti personali relativi a una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti.
Titolare del trattamento	Persona fisica o giuridica, l'autorità pubblica, il servizio o altro organismo che, singolarmente o insieme ad altri, determina le finalità e i mezzi del trattamento di dati personali.
Consenso dell'interessato	Qualsiasi manifestazione di volontà libera, specifica, informata e inequivocabile dell'interessato, con la quale lo stesso manifesta il proprio assenso, mediante dichiarazione o azione positiva inequivocabile, che i dati personali che lo riguardano siano oggetto di trattamento. E' importante distinguere il consenso dell'avente diritto per quanto attiene i profili privacy e quindi il trattamento dei suoi dati personali, dal consenso informato che attiene il contenuto del trattamento medico sanitario al quale il paziente si sottopone. Sono due consensi distinti.

Il GDPR nell'ambito della ricerca medica epidemiologica

Nell'ambito della ricerca medica, epidemiologica, l'Articolo 89 GDPR - Garanzie e deroghe relative al trattamento a fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici recita che *"Il trattamento a fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici è soggetto a garanzie adeguate per i diritti e le libertà dell'interessato, in conformità del presente regolamento. Tali garanzie assicurano che siano state predisposte misure tecniche e organizzative, in particolare al fine di garantire il rispetto del principio della minimizzazione dei dati. Tali misure possono includere la pseudonimizzazione, purché le finalità in questione possano essere conseguite in tal modo. Qualora possano essere conseguite attraverso il trattamento ulteriore che non consenta o non consenta più di identificare l'interessato, tali finalità devono essere conseguite in tal modo."* Pertanto, il trattamento può riguardare unicamente i dati personali strettamente pertinenti gli scopi di ricerca, ivi compresi quelli ricavati da campioni biologici, salvo che questi non costituiscano dati genetici. In particolare, per il trattamento di dati genetici, è necessario fare riferimento all'autorizzazione del Garante n. 8/2016 e, per i trattamenti di dati personali effettuati per scopi di ricerca scientifica, all'autorizzazione n. 9/2016 integrata e aggiornata dal provvedimento del Garante del 5 giugno 2019, rilasciate dal Garante sulla base dell'art. 90 del Codice Privacy (25).

In molti casi non è possibile individuare pienamente la finalità del trattamento dei dati personali a fini di ricerca scientifica al momento della raccolta dei dati. Pertanto, dovrebbe essere consentito agli interessati di prestare il proprio consenso a taluni settori della ricerca scientifica laddove vi sia rispetto delle norme deontologiche riconosciute per la ricerca scientifica. Gli interessati dovrebbero avere la possibilità di prestare il proprio consenso soltanto a determinati settori di ricerca o parti di progetti di ricerca nella misura consentita dalla finalità prevista (Articolo 33 GDPR Notifica di una violazione dei dati personali all'autorità di controllo) (4). È importante che il consenso sia prestato in maniera sufficientemente specifica per settori, ambiti o parti della ricerca così che il paziente – proprietario del dato personale – non incorra in trattamenti che non siano in alcun modo riconducibili all'ambito per il quale il consenso stesso era stato prestato.

Combinando informazioni provenienti dai registri, i ricercatori possono ottenere nuove conoscenze di grande utilità relativamente a patologie diffuse come le malattie cardiovascolari, il cancro e la depressione. Avvalendosi dei registri, i risultati delle ricerche possono acquistare maggiore rilevanza, dal momento che si basano su una popolazione più ampia. I risultati delle ricerche ottenuti dai registri forniscono conoscenze solide e di alta qualità, che possono costituire la base per l'elaborazione e l'attuazione di politiche basate sulla conoscenza, migliorando la qualità della vita per molte persone. Al fine di facilitare la ricerca scientifica, i dati

personali possono essere trattati per finalità di ricerca scientifica fatte salve condizioni e garanzie adeguate previste dal diritto dell'Unione o degli Stati membri (Art 157 del GDPR- Richiesta di informazioni e di esibizione di documenti) (4).

Il termine "ricerca scientifica" dovrebbe includere le diverse tipologie della stessa, ad esempio, sviluppo tecnologico e dimostrazione, ricerca fondamentale e ricerca applicata, finanziata da soggetti pubblici o privati.

Le finalità di ricerca scientifica dovrebbero altresì includere gli studi svolti nell'interesse pubblico nel settore della sanità pubblica (Considerando l'articolo 159 del GDPR) (4).

ANALISI DI CASI REALI

In questo capitolo vengono presentate alcune interazioni, riguardanti casi reali, avvenute fra ricercatori e giuristi, che esemplificano situazioni comuni nell'ambito della medicina di laboratorio, al fine di individuare possibili risposte su elementi specifici.

Esempio 1

Scenario: Studio retrospettivo incentrato sull'analisi di dati relativi ad un esame diagnostico con lo scopo di rendere più efficiente il processo analitico del laboratorio. Nello specifico, lo scopo era di evitare di eseguire un esame di secondo livello cercando di predire il risultato con i soli dati dell'esame di primo livello. Database di circa 10 000 pazienti, dati a disposizione relativi ad età, sesso, risultati dell'esame di primo e secondo livello. Tutti i dati sono stati raccolti nel corso di una routine clinica senza esami aggiuntivi.

Interazione 1. Il ricercatore: e' necessario il consenso al trattamento dei dati?

Il parere del giurista: se i dati sono effettivamente anonimi ovvero senza alcuna possibilità di collegare le informazioni a una persona fisica identificata o identificabile essi non rientrano nell'ambito di applicazione del GDPR Reg. 679/2016 (4). La valutazione del rischio di reidentificazione degli individui attraverso i dati personali elaborati dal laboratorio ed eventuali dati aggiuntivi deve essere effettuata nel caso specifico. I dati personali resi sufficientemente anonimi non consentono più l'identificazione dell'interessato (Figura 4). Se invece i dati fossero pseudonimizzati ovvero consentissero, mediante l'utilizzo di ulteriori informazioni, una identificazione dell'individuo, allora tali dati dovrebbero essere considerati informazioni su una persona fisica identificabile e pertanto rientranti nell'ambito di applicazione del Regolamento 679/2016 per la tutela dei dati personali degli individui. In quest'ultimo caso sarebbe necessario definire la base giuridica per il trattamento: potrebbe essere l'attività di ricerca se vagliata e sostenuta dal DPO della struttura oppure aver acquisito "a monte" un consenso dall'interessato per finalità di ricerca (Figura 5).

Interazione 2. Il ricercatore: lo studio è stato presentato al Comitato Etico della struttura di afferenza del ricercatore, che non lo ha autorizzato

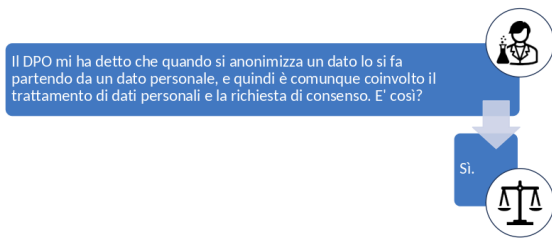


Figura 4

Domanda del ricercatore e risposta del giurista sull'esempio 1: trattamento dei dati personali e anonimizzazione.

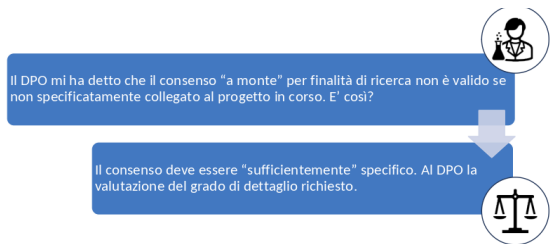


Figura 5

Domanda del ricercatore e risposta del giurista sull'esempio 1: specificità del consenso informato

ritenendo "non chiara la situazione della privacy" e necessario l'ottenimento del consenso informato. Lo studio, quindi, si arresta. Il ricercatore, trasferitosi in altra struttura, prende accordi per condurre comunque lo studio; presenta il progetto all'ufficio studi clinici, che pur riconoscendo l'anonimizzazione dei dati, suggerisce di sottoporre lo studio al parere del Comitato Etico, sebbene per gli studi retrospettivi non vi fosse l'obbligo di approvazione. Infatti, in vista di una possibile pubblicazione scientifica sarebbe stato necessario l'ottenimento dell'approvazione da parte del Comitato Etico, come in passato era accaduto per studi simili.

Il parere del giurista: il passaggio dal Comitato Etico è consigliabile e non implica necessariamente complicazioni per quanto attiene il profilo privacy.

L'epilogo: avendo già perso mesi, il ricercatore decide di proseguire senza il parere del Comitato Etico, pensando di scrivere nella sezione etica dell'articolo che la normativa nazionale non prevede obbligo di approvazione per uno studio retrospettivo su dati anonimi.

Esempio 2

Scenario: Studio retrospettivo incentrato sulla diagnostica differenziale delle patologie dell'emostasi e basato su dati di laboratorio ricavati dal sistema informatico del laboratorio. Nello studio erano coinvolti anche gli ematologi al fine di ottenere le informazioni cliniche sulle coagulopatie di cui erano affetti i pazienti arruolati. Di questi pazienti sarebbero quindi stati recuperati i risultati di esami di laboratorio eseguiti negli anni precedenti. Lo studio rientrava in un progetto la cui attuazione doveva partire entro 6 mesi.

Il ricercatore: il protocollo viene inviato al Comitato Etico, che richiede modifiche sostanziali al documento;

tra queste vi era l'indicazione di rendere chiaro il tipo di studio in quanto non poteva essere considerato di tipo retrospettivo e in questo caso la modulistica permetteva di scegliere lo studio osservazionale trasversale come tipologia più affine. Il Comitato Etico richiede anche di specificare nel dettaglio le motivazioni che rendevano indispensabili i dati di ulteriori esami di laboratorio, motivandone l'utilità, e spiegando quali variabili erano indispensabili per il raggiungimento dell'obiettivo dello studio. Tale spiegazione, evidentemente, non poteva essere correttamente esplicitata, in quanto l'identificazione di eventuali algoritmi (basati su tecniche statistiche tra cui l'analisi supervisionata come indicato nella parte statistica) richiedeva (almeno inizialmente) l'utilizzo di tutte le variabili a disposizione. Inoltre, veniva richiesto di specificare come i dati venivano anonimizzati e il calcolo della potenza statistica, quest'ultima di difficile stima per analisi del tipo proposto.

Il parere del giurista: e' opportuno indicare nel dettaglio come i dati personali del database di laboratorio vengono anonimizzati, cioè la procedura per tale anonimizzazione al fine di poter escludere l'applicazione del Regolamento 679/2016 (Figura 6). In ogni caso, la valutazione del rischio di reidentificazione degli individui attraverso i dati personali elaborati dal laboratorio ed eventuali dati aggiuntivi deve essere effettuata all'interno del caso concreto.

L'epilogo: dato il tempo esiguo a disposizione, il protocollo revisionato è inviato molto rapidamente. Il Comitato Etico richiede ulteriori importanti modifiche al protocollo, ma il tempo era ormai terminato.

Esempio 3

Scenario: Studio retrospettivo incentrato su dati di laboratorio provenienti da screening regionale effettuato su minori (neonati) per migliorare la capacità di selezione e il percorso di cura. Lo studio prevedeva la descrizione della popolazione, la definizione di valori decisionali, di valori di riferimento. Servivano, inoltre, i dati clinici (positività/negatività/rischio) relativi alle malattie oggetto dello screening.

Interazione 1. Il ricercatore: per i soggetti positivi allo screening, non dovrebbe esserci alcun problema perché sono presi in carico dal centro di cura e per questo firmano un consenso che abilita anche la ricerca (Figura 7). I soggetti negativi allo screening invece escono dal programma: come utilizzare i dati dei negativi? Serve il consenso? Come ottenerlo?

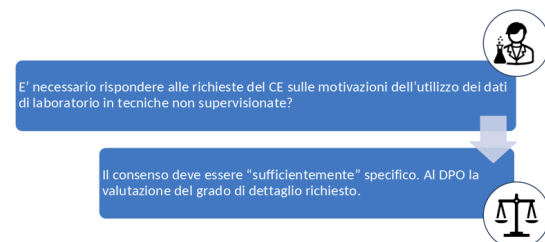


Figura 6

Domanda del ricercatore e risposta del giurista sull'esempio 2: motivazioni dell'utilizzo dei dati.

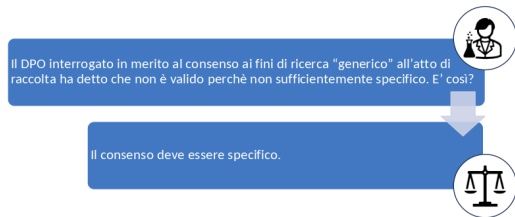


Figura 7

Domanda del ricercatore e risposta del giurista sull'esempio 3: consenso "generico" ai fini di ricerca

Il parere del giurista: se sono dati anonimi non serve consenso, se invece non lo sono, è necessario il consenso al trattamento che dovrebbe essere effettuato "a monte" della procedura. È infatti ideale che il consenso venga preso all'avvio di qualsiasi analisi/screening così da includere fin da subito eventuali trattamenti di dati personali, afferenti allo stato di salute per i quali sia eventualmente necessario il consenso del paziente. Si ricorda la distinzione tra consenso al trattamento dei dati e consenso informato alle cure/trattamenti medici. Inoltre, si ricorda che il consenso al trattamento dei dati personali deve essere specifico e sufficientemente dettagliato, tale da fornire al paziente tutte le informazioni sulle finalità del trattamento di utilizzo dei suoi dati.

Interazione 2. Il ricercatore: supponendo che l'approvazione del Comitato Etico arrivi sulla base di una descrizione della popolazione sottoposta ad analisi di screening per sindromi metaboliche, e che analizzando i dati si ritenga di approfondire la correlazione con patologie (ad esempio, le emoglobinopatie) non comprese nella definizione "malattie metaboliche" del consenso originario, come bisogna comportarsi?

Il parere del giurista: le finalità di ricerca possono essere la base giuridica per tutte le patologie, si consiglia di indicarle in maniera ampia e comunque di includere la dicitura "patologie metaboliche o in qualsiasi modo patologie ad esse correlate".

Interazione 3. Il ricercatore: un ulteriore punto molto importante riguarda l'analisi dei Big Data, effettuata per esempio con tecniche di apprendimento non supervisionato, che potrebbe permettere di vedere clusterizzazioni del tutto nuove, che non sono ancora state identificate e che quindi non erano state previste. E ancora, nel caso di dati insoliti, che rendono potenzialmente possibile l'identificazione della persona anche senza anagrafica (ad esempio: madre di 10 figli), come è possibile procedere per gestire e pubblicare il dato?

Il parere del giurista: se il dato consente una identificazione dell'interessato è necessario avere una base giuridica per il suo trattamento: o il dato pseudonimizzato (madre 10 figli comunque non riporta il nome) è trattato sulla base dell'eccezione della ricerca scientifica o sulla base del consenso del paziente.

L'epilogo: Il progetto è fermo: dopo un anno non è stato ancora sottoposto al Comitato Etico in attesa dell'esito dell'interrogazione al Garante effettuata dal DPO.

Esempio 4

Scenario: Studio retrospettivo multicentrico no-profit mirato ad ottenere una stima dell'incidenza annuale di una patologia congenita di interesse, utilizzando spot di sangue raccolto a 48-72 ore di vita per lo screening neonatale presso i Punti Nascita regionali. La rilevanza dei dati raccolti è perlopiù epidemiologica.

Il ricercatore: Lo studio retrospettivo prevede l'utilizzo dello spot di sangue in forma anonima e non prevede la raccolta di dati personali. Ai fini dello studio non è infatti necessario acquisire informazioni quali età, sesso e così via. Il protocollo di studio può essere autorizzato dal Comitato Etico senza informativa e consenso dei genitori?

Nella Tabella 6 si riporta la lista della documentazione richiesta da Comitato Etico di riferimento per gli studi osservazionali.

Il parere del giurista: Se non sono coinvolti dati personali, il consenso al trattamento dei dati personali non è rilevante perché non si rientra nell'ambito di applicazione del Regolamento 679/2016. Si sottolinea, inoltre, che ai fini privacy è irrilevante se lo studio sia no-profit.

L'epilogo: lo studio è fermo: sono presenti dati personali e il consenso non si può ottenere retrospettivamente.

Tabella 6

Documentazione richiesta dal Comitato Etico di riferimento del caso 4 per gli studi osservazionali

Esempio di elenco di documenti richiesti da un Comitato Etico per gli studi osservazionali

- LETTERA DI TRASMISSIONE indirizzata a Direttore Generale, CEI e sperimentatore principale
- LISTA DELLE AUTORITA' COMPETENTI (di altri paesi a cui è stata presentata la domanda e relative decisioni (se disponibili))
- DELEGA O LETTERA DI INCARICO che autorizzi a presentare domanda di autorizzazione in nome e per conto del promotore/sponsor (qualora il richiedente non sia il promotore stesso)
- CURRICULUM VITAE del ricercatore principale
- DICHIARAZIONE DI ASSENZA DI CONFLITTO DI INTERESSI del ricercatore principale
- DICHIARAZIONE SULLA NATURA OSSERVAZIONALE DELLO STUDIO
- PARERE UNICO CE del centro coordinatore (per domande presentate da centri collaboratori)
- ELENCO DEI CENTRI PARTECIPANTI e relativi responsabili
- PROTOCOLLO DELLO STUDIO
- SINOSI in italiano (seguire il modello di compilazione)
- SCHEDA RACCOLTA DATI
- FOGLIO INFORMATIVO per il paziente (da redigere su carta intestata)
- MODULO DI CONSENSO INFORMATO (da redigere su carta intestata)
- LETTERA INFORMATIVA PER IL MEDICO CURANTE
- APPROVAZIONE DS (Direzione Sanitaria)
- MODULO RSO
- BOZZA DI CONVENZIONE ECONOMICA promotore e sede di svolgimento dello studio
- ALTRI DOCUMENTI utili di qualunque tipo
- ELENCO DELLA DOCUMENTAZIONE PRESENTATA, con indicazione, per ognuno, del numero della versione e della data

CONFLITTO DI INTERESSE

Nessuno

BIBLIOGRAFIA

- Blatter TU, Witte H, Nakas CT, Leichtle AB. Big Data in Laboratory Medicine-FAIR Quality for AI? *Diagnostics (Basel)*. 2022;12:1923.
- Mannello F, Plebani M. Current issues, challenges, and future perspectives in clinical laboratory medicine. *J Clin Med* 2022;11:634.
- Ronzio L, Cabitza F, Barbaro A, Banfi G. Has the flood entered the basement? A systematic literature review about machine learning in laboratory medicine. *Diagnostics (Basel)* 2021;11:372.
- Regolamento generale sulla protezione dei dati - (UE) 2016/679. <https://www.garanteprivacy.it/documenti/10160/0/Regolamento+UE+2016+679.+Arricchito+con+riferimenti+ai+Considerando+Aggiornato+alle+rettifiche+pubblicate+sulla+Gazzetta+Ufficiale+dell'Unione+europea+127+del+23+maggio+2018> (ultimo accesso: novembre 2023).
- Tribunale dell'Unione (causa T-557/20, sentenza 26 aprile 2023). <https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:62020TJ0557> (ultimo accesso: novembre 2023).
- Attività dei comitati etici istituiti ai sensi del decreto ministeriale 18 marzo 1998. GU Serie Generale n.214 del 12-09-2002. https://www.gazzettaufficiale.it/atto/serie_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=2002-09-12&atto.codiceRedazionale=02A11274&elenco30giorni=false (ultimo accesso: novembre 2023).
- Linee guida per la classificazione e conduzione degli studi osservazionali sui farmaci. <https://www.aifa.gov.it/documents/20142/516919/111.88758.1186138046156a-0be.pdf> (ultimo accesso: novembre 2023).
- Misure volte a facilitare e sostenere la realizzazione degli studi clinici di medicinali senza scopo di lucro e degli studi osservazionali e a disciplinare la cessione di dati e risultati di sperimentazioni senza scopo di lucro a fini registrativi. <https://www.gazzettaufficiale.it/eli/id/2022/02/19/22A01189/sg> (ultimo accesso: novembre 2023).
- Decreto legislativo 10 agosto 2018, n. 101. <https://www.garanteprivacy.it/documents/10160/0/Codice+in+materia+di+protezione+dei+dati+personali+%28Testo+coordinato%29> (ultimo accesso: novembre 2023).
- Criticità etiche e normative nel trattamento dei dati personali sanitari nella ricerca osservazionale. https://www.aifa.gov.it/documents/20142/1808580/Criticita_etiche_ricerca_osservazionale_06.04.2023.pdf (ultimo accesso: novembre 2023).
- Bellini C, Padoan A, Carobene A, Guerranti R. A survey on Artificial Intelligence and Big Data utilisation in Italian clinical laboratories. *Clin Chem Lab Med* 2022;60:2017-26.
- Lippi G, Plebani M. Integrated diagnostics: the future of laboratory medicine? *Biochem Med* 2020;30:e010501.
- Pisapia A, Banfi G, Tomaiuolo R. The novelties of the regulation on health technology assessment, a key achievement for the European union health policies. *Clin Chem Lab Med* 2022;60:1160-3.
- Khan S, Khan HU, Nazir S. Systematic analysis of healthcare big data analytics for efficient care and disease diagnosing. *Sci Rep* 2022;12:1-21.
- Gruson D. Big data, artificial intelligence and laboratory medicine: time for integration. *Adv Lab Med* 2021;2:1-3.
- Cabitza F, Banfi G. Machine learning in laboratory medicine: Waiting for the flood? *Clin Chem Lab Med* 2018;56:516-24.
- Zuo Z, Watson M, Budgen D, Hall R, Kennelly C, Al Moubayed N. Data Anonymization for Pervasive Health Care: Systematic Literature Mapping Study. *JMIR Med Inform* 2021;9:e29871.
- El Emam K, Rodgers S, Malin B. Anonymising and sharing individual patient data. *Br Med J* 2015;350:h1139.
- Vijayarani S, Tamilarasi A. An efficient masking technique for sensitive data protection. Proceedings of the IEEE 2011 International Conference on Recent Trends in Information Technology. Chennai, India 2011:03-05. <https://ieeexplore.ieee.org/document/5972275> (ultimo accesso: novembre 2023).
- Kelly JP, Golden BL, Assad AA. Cell suppression: disclosure protection for sensitive tabular data. *Netw Int J* 1992;22:397-417.
- Samarati P, Sweeney L. Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression. Technical Report, Computer Science Laboratory, SRI International. 1998.
- Machanavajhala A, Gehrke J, Kifer D, Venkatasubramanian M. L-diversity: privacy beyond k-anonymity. Proceedings of the IEEE 22nd International Conference on Data Engineering. 2006, pp. 03-07.
- Li N, Li T, Venkatasubramanian S. t-Closeness: Privacy Beyond k-Anonymity and l-Diversity. Proceedings of the IEEE 23rd International Conference on Data Engineering. 2007, pp. 106-115.
- Soria-Comas J, Domingo-Ferrer J, Sánchez D, Martínez S. t-closeness through microaggregation: Strict privacy with enhanced utility preservation. Proceedings of the IEEE 32nd International Conference on Data Engineering. 2016, pp. 1464-1465.
- Provvedimento che individua le prescrizioni contenute nelle Autorizzazioni generali nn. 1/2016, 3/2016, 6/2016, 8/2016 e 9/2016 che risultano compatibili con il Regolamento e con il d.lgs. n. 101/2018 di adeguamento del Codice - 13 dicembre 2018 (9068972). <https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9068972> (ultimo accesso: novembre 2023).